

Nigie

PCT

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/830,691

DATE: 05/18/2001
TIME: 10:56:06

Input Set : A:\118.12-US-WO SEQLIST.txt
Output Set : N:\CRF3\05182001\I830691.raw

ENTERED

4 <110> APPLICANT: Choi, Eui-Sung
5 Rhee, Sang-Ki
6 Sohn, Jung-Hoon
7 Park, Soo-Dong
8 Lee, Yoon-Hyoung
9 Lee, Seung-Jae
10 Jang, Jae-Kweon
11 Choi, Seok-Keun
12 Son, Young-Rok
14 <120> TITLE OF INVENTION: VECTOR FOR THE TRANSFORMATION OF PHAFFIA
15 RHODOZYMA AND PROCESS OF TRANSFORMATION THEREBY
18 <130> FILE REFERENCE: 118.12-US-WO
C--> 20 <140> CURRENT APPLICATION NUMBER: US/09/830,691
C--> 21 <141> CURRENT FILING DATE: 2001-04-26
22 <150> PRIOR APPLICATION NUMBER: KR 1998/46547
23 <151> PRIOR FILING DATE: 1998-10-31
24 <160> NUMBER OF SEQ ID NOS: 14
25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1223
28 <212> TYPE: DNA
29 <213> ORGANISM: Phaffia rhodozyma
30 <400> SEQUENCE: 1
31
32
33
34
35
36 atggtcaacg ttcccaagac tcgacgtgag ttatagcaat ttcaacaact ctccagacga 60
37 caaatattcc agtgcacga aagagtttgt ggataaacgc gacagtttca agggaaagag 120
38 tcgatggaca gatttggag acttagccgg tcaaggaaact tggggatcac gtggcggagg 180
39 actcatcaga agaagtcggg atttgtttga tcatagtggg atcaagacaa actggaggat 240
40 atggctcgcc ttggaaggga atctccggcc tggattcgag gatccgaaag ttgtacgtat 300
41 ggaaaagctt acacggcttg gatttattat ctttcatagg aacctactgc aagggttaag 360
42 cttgcaagaa gcacacgtaa gtcgcttatt ctctccactc tttcatggca tattgtcaac 420
43 gactggacaa cgcgtccggt ttgaaacaag tgacttacct gtgaaatttg attctacacc 480
44 tgtatttagc cctcacaagg tacatatcac atcctccac cccaccctgc ccaacttctt 540
45 cagttcatct tgcctcctgt ttccacattc cctgatgacc tccttgtagg ttctttgcga 600
46 acgtttgttt ctgtttctgt aggtgaccca gtacaagaag ggaaaggact ccatcttcgc 660
47 ccagggaaag cgacgatacg accgaaagca gtccgggttac ggaggtcaga ccaagcccgt 720
48 tttccacaag aaggctaaga ccaccaagaa ggtcgtcctt cgattggcgg tatttttgtt 780
49 tattttgaat tctttttgtg tatgcagact tttgatgatt atgctcctct gtcgtttttt 840
50 ctcttcaaac agagtgtctc gtctgcagtt cgttcttctt tccaaccaa acttcaacta 900
51 cagacatcat aaacagacat cttacttcgg tgttctctct tttttccgc agagtacaag 960
52 atgcagatga cctcgaagcg atgcaagcac ttcgagcttg gaggagacaa gaagaccaag 1020
53 ggttcgtctt ttgtccatat attctctggt tcaactctta tgttctaac gtacttgttt 1080
54 ccttttttgt tcggatgttg tttctatcgg tgggtttttc ttttctttgg atgcattatc 1140
55 atttatcgtg ttggactggt ttctctgctt cgtttctttc tctctgttac ttgtgcttct 1200
56 caggagccgc catctctttc taa 1223
58 <210> SEQ ID NO: 2
59 <211> LENGTH: 350
60 <212> TYPE: DNA

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,691

DATE: 05/18/2001

TIME: 10:56:06

Input Set : A:\118.12-US-WO SEQLIST.txt

Output Set: N:\CRF3\05182001\I830691.raw

```

61 <213> ORGANISM: Phaffia rhodozyma
63 <220> FEATURE:
64 <221> NAME/KEY: CDS
65 <222> LOCATION: (30)...(347)
67 <400> SEQUENCE: 2
68 cccttcaagt ctgctctcaa tcagtcaag atg gtc aac gtt ccc aag act cga      53
69                               Met Val Asn Val Pro Lys Thr Arg
70                               1           5
72 cga acc tac tgc aag ggt aag gct tgc aag aag cac acc cct cac aag      101
73 Arg Thr Tyr Cys Lys Gly Lys Ala Cys Lys Lys His Thr Pro His Lys
74   10           15           20
76 gtg acc cag tac aag aag gga aag gac tcc atc ttc gcc cag gga aag      149
77 Val Thr Gln Tyr Lys Lys Gly Lys Asp Ser Ile Phe Ala Gln Gly Lys
78   25           30           35           40
80 cga cga tac gac cga aag cag tcc ggt tac gga ggt cag acc aag ccc      197
81 Arg Arg Tyr Asp Arg Lys Gln Ser Gly Tyr Gly Gly Gln Thr Lys Pro
82           45           50           55
84 gtt ttc cac aag aag gct aag acc acc aag aag gtc gtc ctt cga ttg      245
85 Val Phe His Lys Lys Ala Lys Thr Thr Lys Lys Val Val Leu Arg Leu
86           60           65           70
88 gag tgc tcc gtc tgc aag tac aag atg cag atg acc ctc aag cga tgc      293
89 Glu Cys Ser Val Cys Lys Tyr Lys Met Gln Met Thr Leu Lys Arg Cys
90           75           80           85
92 aag cac ttc gag ctt gga gga gac aag aag acc aag gga gcc gcc atc      341
93 Lys His Phe Glu Leu Gly Gly Asp Lys Lys Thr Lys Gly Ala Ala Ile
94           90           95           100
96 tct ttc taa      350
97 Ser Phe
98 105
101 <210> SEQ ID NO: 3
102 <211> LENGTH: 106
103 <212> TYPE: PRT
104 <213> ORGANISM: Phaffia rhodozyma
106 <400> SEQUENCE: 3
107 Met Val Asn Val Pro Lys Thr Arg Arg Thr Tyr Cys Lys Gly Lys Ala
108   1           5           10           15
109 Cys Lys Lys His Thr Pro His Lys Val Thr Gln Tyr Lys Lys Gly Lys
110           20           25           30
111 Asp Ser Ile Phe Ala Gln Gly Lys Arg Arg Tyr Asp Arg Lys Gln Ser
112           35           40           45
113 Gly Tyr Gly Gly Gln Thr Lys Pro Val Phe His Lys Lys Ala Lys Thr
114           50           55           60
115 Thr Lys Lys Val Val Leu Arg Leu Glu Cys Ser Val Cys Lys Tyr Lys
116   65           70           75           80
117 Met Gln Met Thr Leu Lys Arg Cys Lys His Phe Glu Leu Gly Gly Asp
118           85           90           95
119 Lys Lys Thr Lys Gly Ala Ala Ile Ser Phe
120           100           105
122 <210> SEQ ID NO: 4

```

RAW SEQUENCE LISTING

DATE: 05/18/2001

PATENT APPLICATION: US/09/830,691

TIME: 10:56:06

Input Set : A:\118.12-US-WO SEQLIST.txt

Output Set: N:\CRF3\05182001\I830691.raw

```

123 <211> LENGTH: 741
124 <212> TYPE: DNA
125 <213> ORGANISM: Phaffia rhodozyma
127 <220> FEATURE:
128 <221> NAME/KEY: misc_feature
129 <222> LOCATION: (0)...(0)
130 <223> OTHER INFORMATION: n=a, t, c, or g
132 <400> SEQUENCE: 4
133 ctcgagtggg cgggtggcaat ggcattcgtg tcgttggtgc tcaactcgcaa cccaagcagt      60
134 cgcttaccgc gggtagcctc cgggtgggcg cgatgatttg tgggtgggat tccttcctta      120
135 tgggtagaac gacgcgcaac caatcattcg gagaaccgct ccgttgtagc cgaccagtct      180
136 gattgatcaa catgccagca cgtcctcccg gacggagact ggcggggatc gtacctcatc      240
137 tggaatcgct ggctcaatgg tagtagcttt cacgatcggc catgagggca gtctaggtgg      300
W--> 138 gttcgcctgc cgaagactgt gtgagtgtgc tganaactaa ttgagtaccg ggggataagg      360
W--> 139 caaggcgtgt ntggttgccg gtggctgtga gcgagtttgc tgcaaagcga ttcaatgcac      420
140 cccggccttg ccagcgcgct gcgtcacgaa acacactaaa cggttgacgc cataaagtaa      480
141 taacacactc aagtttgttg tcccgggtgg gcctctgtgc ctgcgtggga cccgacggga      540
142 gaggaaaacg ttctgtggcc ctctcctctg tggatagtta cctggttgat cctgccagta      600
143 gtcatatget tgtctcaaag attaaagccat gcatgtctaa gtataaacia attcatactg      660
144 tgaactgcgc aatggctcat taaatcagtt atagttttatt tgatggtacc ttgctacatg      720
145 gataactgtg gtaattctag a                                     741
147 <210> SEQ ID NO: 5
148 <211> LENGTH: 23
149 <212> TYPE: DNA
150 <213> ORGANISM: Artificial Sequence
152 <220> FEATURE:
153 <223> OTHER INFORMATION: CYH1, a PCR primer for the cloning of L41 genomic
154 DNA fragment
156 <221> NAME/KEY: misc_feature
157 <222> LOCATION: (0)...(0)
158 <223> OTHER INFORMATION: n=a, t, c, or g
160 <400> SEQUENCE: 5
W--> 161 cgcgtagtta aygtncnaa rac                                     23
163 <210> SEQ ID NO: 6
164 <211> LENGTH: 25
165 <212> TYPE: DNA
166 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: CYH3, a PCR primer for the cloning of L41 genomic
170 DNA fragment
172 <400> SEQUENCE: 6
173 cccgggtytt ggcyttyttr tgraa                                     25
175 <210> SEQ ID NO: 7
176 <211> LENGTH: 24
177 <212> TYPE: DNA
178 <213> ORGANISM: Artificial Sequence
180 <220> FEATURE:
181 <223> OTHER INFORMATION: 3' RACE primer
183 <400> SEQUENCE: 7

```

RAW SEQUENCE LISTING

DATE: 05/18/2001

PATENT APPLICATION: US/09/830,691

TIME: 10:56:06

Input Set : A:\118.12-US-WO SEQLIST.txt

Output Set: N:\CRF3\05182001\I830691.raw

```

184 ggtcagacca agcaagtttt tcac 24
186 <210> SEQ ID NO: 8
187 <211> LENGTH: 24
188 <212> TYPE: DNA
189 <213> ORGANISM: Artificial Sequence
191 <220> FEATURE:
192 <223> OTHER INFORMATION: 5' RACE primer
194 <400> SEQUENCE: 8
195 gtgaaaaaact tgcttggtct gacc 24
197 <210> SEQ ID NO: 9
198 <211> LENGTH: 24
199 <212> TYPE: DNA
200 <213> ORGANISM: Artificial Sequence
202 <220> FEATURE:
203 <223> OTHER INFORMATION: sense primer for the mutagenesis of L41 gene
205 <400> SEQUENCE: 9
206 ggtcagacca agcaagtttt tcac 24
208 <210> SEQ ID NO: 10
209 <211> LENGTH: 24
210 <212> TYPE: DNA
211 <213> ORGANISM: Artificial Sequence
213 <220> FEATURE:
214 <223> OTHER INFORMATION: antisense primer for the mutagenesis of L41 gene
216 <400> SEQUENCE: 10
217 gtgaaaaaact tgcttggtct gacc 24
219 <210> SEQ ID NO: 11
220 <211> LENGTH: 20
221 <212> TYPE: DNA
222 <213> ORGANISM: Artificial Sequence
224 <220> FEATURE:
225 <223> OTHER INFORMATION: a PCR primer corresponding to 18S rDNA
227 <400> SEQUENCE: 11
228 tcctagtaag cgcaagtcac 20
230 <210> SEQ ID NO: 12
231 <211> LENGTH: 20
232 <212> TYPE: DNA
233 <213> ORGANISM: Artificial Sequence
235 <220> FEATURE:
236 <223> OTHER INFORMATION: a PCR primer corresponding to 18S rDNA
238 <400> SEQUENCE: 12
239 ttcggccaag gaaagaaaact 20
241 <210> SEQ ID NO: 13
242 <211> LENGTH: 20
243 <212> TYPE: DNA
244 <213> ORGANISM: Artificial Sequence
246 <220> FEATURE:
247 <223> OTHER INFORMATION: a PCR primer corresponding to 28S rDNA
249 <400> SEQUENCE: 13
250 aatcggatta tccggagcta 20

```

DATE: 05/18/2001

TIME: 10:56:06

Input Set : A:\118.12-US-WO SEQLIST.txt

Output Set: N:\CRF3\05182001\I830691.raw

252 <210> SEQ ID NO: 14

253 <211> LENGTH: 20

254 <212> TYPE: DNA

255 <213> ORGANISM: Artificial Sequence

257 <220> FEATURE:

258 <223> OTHER INFORMATION: a PCR primer corresponding to 28S rDNA

260 <400> SEQUENCE: 14

261 gctataacac atccggagat

20

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

DATE: 05/18/2001

TIME: 10:56:07

Input Set : A:\118.12-US-WO SEQLIST.txt

Output Set: N:\CRF3\05182001\I830691.raw

L:20 M:270 C: Current Application Number differs, Replaced Current Application Number

L:21 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:138 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:139 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5

[illegible]